Paper: jc*--**-****

Chaotic Music Generation System Using Music Conductor Gesture

Shuai Chen, Yoichiro Maeda, Yasutake Takahashi

Dept. of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui
3-9-1, Bunkyo, Fukui, 910-8507 Japan
E-mail: (schen, maeda, yasutake)@ir.his.u-fukui.ac.jp
[Received 00/00/00; accepted 00/00/00]

Abstract. In the research of interactive music generation, we propose a music generation method, that the computer generates the music, under the recognition of human music conductor's gestures. In this research, the generated music is tuned by the recognized gestures for the parameters of the network of chaotic elements in real time. The music conductor's hand motions are detected by Microsoft Kinect in this system. Music theories are embedded in the algorithm, as a result, the generated music will be richer. Furthermore, we constructed the music generation system and performed the experiment for generating the music composed by human beings.

Keywords: Chaotic Theory, Network of Chaotic Elements, Artificial Sound, Music Generation

1. Introduction

Traditional music has thousands of years of history. But playing a musical instrument is not a simple matter for every person. Only musicians that have undergone a process of intensive training can produce music through an instrument effectively. The people who only have interest but know nothing about the instruments, can only play the role of the listeners. However, interactive music, that typically involves human interaction, has become a new way to create the music.

There are already existing interactive systems controlling music or sound based on human body movements, for example, Sound Sculpting[1]. It proposed a new approach to mapping hand movements to sound through shapes of a virtual input device controlled by both hands. The attributes of the virtual object are translated into parameters for real-time sound editing within MAX/MSP. And a vision-based system, Body-Brush, has been proposed, that captures the entire human body motions and gestures for 3D painting synthesis and musical sound generation[2]. And also, instead of the virtual acquisition of body movements, Cyber Composer, introduced as a music generation system, that melody flow and musical expressions can be controlled and generated by wearing motion-sensing gloves[3].

In order to arouse the interest from everyone, it is necessary to make the users control more in the way to generate their music. In addition, in order to ensure that the generated music makes every user feel easy to understand and easy to control, music conductor gestures must be applied in this paper.

In this research, the proposed method is designed, not only following the requirements of people but also generating the music automatically by using Interactive Chaotic Amusement System (ICAS)[4]. Music conductor gestures and basic music theories are embedded in the algorithm, so that, after the music have been generated, it will be arranged by the human's hand conducting gestures. The users use their hand gestures to interact with the computer for generating the music. Human intention and chaotic calculation generate sounds with mutual multiplier effects.

2. Overview of ICAS

Network of chaotic elements has the connected elements of chaos which are coupled to mainly network-like. And these chaotic elements are given in the form of differential equation. It is also known as large-scale coupled map, proposed by K.Kaneko[5]. By using this theory, it is possible to reduce a complex and various behavior of the entire map. In addition, the combination structure of network of chaotic elements can be divided into Coupled Map Lattice (CML), and Globally Coupled Map (GCM). In this research, the Globally Coupled Map has been used in ICAS.

2.1. Globally Coupled Map

Globally Coupled Map(GCM) is a model of non-linear system with a global connected chaotic network, that changed by all other elements interacting with the same degree of intensity. The equation of GCM is shown as follows:

$$f(x_j(t)) = 1 - ax_i^2(t) \qquad [i = 1, 2, ..., n], \quad . \quad . \quad . \quad (1)$$

$$x_i(t+1) = [1-e]\, f(x_i(t)) + \frac{e}{N}\sum_{j=1}^{N} f(x_j(t)), \quad . \quad (2)$$

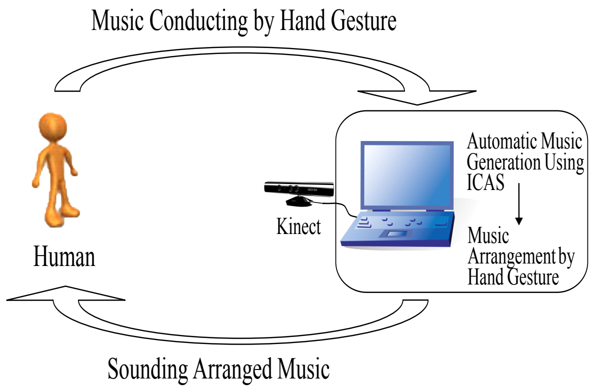


Fig. 1. Overview of Music Generation System

where $x_i(t)$ shows the state of element i at a discrete time t, $f(x)$ shows logistic map, a parameter e shows the strength of the entire combination, and N shows the total number of elements.

2.2. About ICAS

In this research, we propose some further works of Interactive Chaotic Amusement System (ICAS)[4]. By using ICAS which united the chaotic elements to generate various sounds by GCM, favorite sounds for an operator appear as a whole adjusting the parameter of GCM. Its control is enabled by the human operator, and the sound is tuned by a visual information.

3. Proposed Music Generation System

We propose a method to generate the music that changes according to the hand gesture of operator in real-time, and not simply mapping the gestures directly into music. Through the application and extension of ICAS, we make it possible to combine the diversity and the randomness of computer-generated music with the human requirements for the music. The overview of our system is shown in Fig. 1. Fig. 2 shows the algorithm of the proposed system.

In this research, first we get the GCM parameters from the Graphic User Interface(GUI) with these data, and we can generate the pitch and length of sound. And then, with the selected rhythm type in the GUI, we adjust and accompanied the sound with the rhythm and chord progression, and we achieve to make our output sounds more like music. We capture the operator's hand gesture by Kinect, and find out the hand position from the input data, then figure out the center point of hand and wrist as the characteristic point of hand gesture. With the coordinates of those points of both hands in several serial frames, we can calculate the amplitude, speed and acceleration of the motion of hand. Finally, we arranged the automatic generated music with the data of recognized hand gesture.

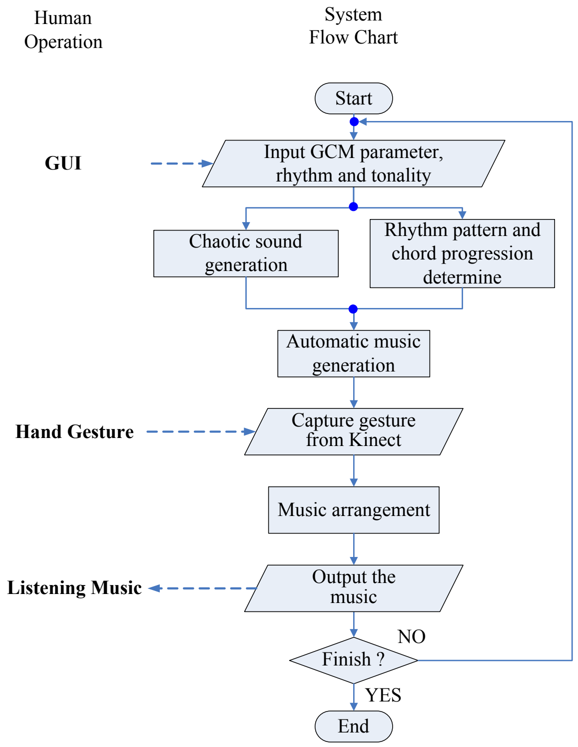


Fig. 2. Algorithm of Music Generation System

4. Music Generation

The ordinary ICAS system united the chaotic elements to generate sound by the Network of Chaotic Elements, and its control is enable by the operator. Through musical styles that involve tonal harmony, chord progression determines the shape of a composition to a great extent, and the harmonic rhythm typically has the function of defining or confirming the prevailing meter of a composition[6]. In our music generation system, we use ICAS to generate the basic sound according to the operator's setting of rhythm and tonality, and we proposed a method to select the chord progression automatically. After this, we obtain the basic sound, generated by ICAS, the rhythm and chord progression, generated by our automatic selection method. Next we synchronize these parts together, to make sure each part starts and ends at the same time. We have four different sound tracks, which all generated individually, and we can let this four track play on various sorts of instruments. Here we determine the first two tracks as the basic sound tracks, the third tracks is for the chord progression, and the fouth track as the rhythm playing track.

4.1. Sound Generation by ICAS

Fig. 3 gives the basic sound generation idea by using ICAS. Unlike the ordinary ICAS, which generates the pitch, length and volume all by GCM, here we only use the GCM to generate the pitch and length. The volume will be controlled by hand gesture directly. The generated sounds that from Note 1 to Note 4 are output simultaneously. While each note is generated respectively, sound i (i=1,2) is produced under the logistic map of Lh_i,

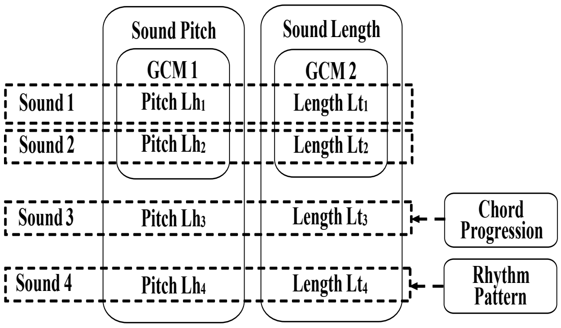


Fig. 3. Sound Generation by ICAS

Lt_i. In addition, the pitch of Lh_1 and Lh_2 have been coupled within a global map GCM1. Similarly, the length are coupled with the global map GCM2. The output of sound generation $x_i(t)$ according to Equation (2), is within $-1.0 < x_i(t) < 1.0$, and this range is divided as follows.

1 Tuning of pitch (GCM1)
We divided this range into 20, thus each small range covers 0.1. These ranges corresponding to the pitch in MIDI, the exact pitch of these 20 ranges will be changed if the selected tonality changes. When the tonality is C major, these are from C−, one octave lower than center C, to A+, two octave higher.

2 Tuning of sound length (GCM2)
We define the longest length as a whole note, which lasts 2000(msec), as a result, we divided the range into 1:2:4:2:1, corresponding to the length of 125:250:500:1000:2000(msec), as the quarter note have the highest probability of occurrence. Hence, if we take the quarter note as one beat, then these sounds will be played at the speed of 120 notes per minutes.

4.2. Rhythm Patterns

We take the rhythm into account, in our music generation system. Rhythmic phrases and patterns have been passing on different approaches in different music style throughout the world. Since the purpose of us is to try to add the rhythm patterns into ICAS generation, and intend to have a better output. In order to confirm our ideas, we come up with a simple implementation of rhythm. In the present stage of the generation experiment, we only added three different rhythms: tango, rumba and samba. For each of these rhythm phrases, there are several typical patterns have been found[7]. Fig. 4 shows the typical rhythm patterns of tango, rumba and samba, which have been utilized in this research.

4.3. Chord Progression

The chord progression is actually a series of musical chords, this series of musical chords can create the harmony. In western classical notation, Roman numerals are used to number the chords built on the scale. For example, I is indicated to the first note of a key, and a V in the key of C is indicated to G chord, but in the key of A it indicated to E chord. They usually occur on an accented beat, and with the rhythm goes on, the chord shifts from one to another. These chord shifts are undergoing some regular patterns, here we used the common three chords progression: I - IV - V, for key of C, this chord progression turn out to be Chord C, Chord F and Chord G. All the chord progression assembled by these three chords, will be always sounds harmony. We randomly select one chord from the three chords, and make it occurs at each accented beat of the rhythm.

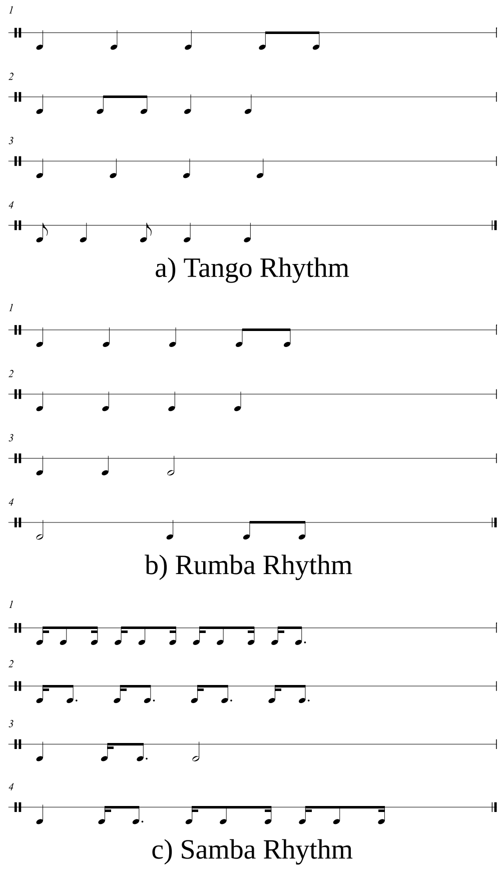


Fig. 4. Typical Rhythm Patterns

4.4. Synchronization

Synchronization is very important in our system, although we generate the four tracks in the same time scale, because any little delay or dislocate between two tracks will cause the output sounds great disarray. After we recognize this problem in the experiments, we tried to solve it by adjust the error of note length generation. The longest generated note lasts 2000(msec), if we let the quarter note as one beat, and we can get the smallest section with four beats in it. When every section finished, we eliminate the error between four tracks once. If the tonality or the rhythm have been changed, then all the tracks reset back into zero, and start the generation from the beginning again.

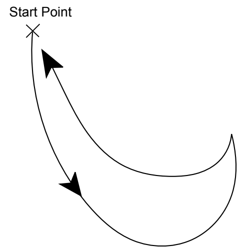


Fig. 5. Hand Tracing Shape of 2/2 or 2/4 Time

5. Hand Gesture Recognition Based on Technique of Music Conducting

Conducting is a means of communicating artistic directions to performers during a performance. Although it is almost subjective and a variety of different conducting styles exists depending upon the training and sophistication of the conductor, there are many formal rules on how to conduct correctly. Despite a wide variety of styles, a number of standard conventions have developed.

5.1. Hand Gesture for Beat and Tempo

The beat of the music, is usually expressed by the conductor's right hand. The shape of hand is waving in the air at every bar (a short section of music) depending on the time signature of certain piece of music, every change from the downward to upward of hand movement is called one beat. Fig. 5 shows the simplest common beat patterns considering the 2/2 or 2/4 time. At this stage, we used this pattern only.

The length of each beat, the time that the hand moves downward to upward once, determines the length of the note, and the speed of the music. Music conductor usually changes the wave speed of his/her hands to transmit to musicians, the music here should be played in a different speed.

Therefore, in our system, we calculate the the time of hand downward to upward, and take the speed of the characteristic point of hand as a characteristic value of sound length. We change the tempo from 60 Beats Per Minute (BPM) to 120 BPM, according to the range of moving speed of hand characteristic point from 300 pixels per second to 3000 pixels per second.

5.2. Hand Gesture for Dynamics

The music dynamics is indicated by the amplitude of conductor's hand movement, the larger shape expresses the louder sound. The conductors usually use their left hand, while right hand have been used to show the beat, to evince the extent of dynamic. When the hand moves upward (usually palm-up) represents a crescendo, and when the hand moves downward (usually palm-down) means a diminuendo. Changing the amplitude of hand movement depending on the circumstances, often leads the changing of the characteristic of the music. Here we try to make it simple, so we use the amplitude of the hand as a characteristic value of volume. If the conductor wave his hands of a great range in a beat, the output music volume will be great, and if the hand waving is smaller, output volume will be turned down. Here we mapped the range of amplitude of hand characteristic point from 20 pixels to 500 pixels, to the MIDI volume in the range of 0 to 127.

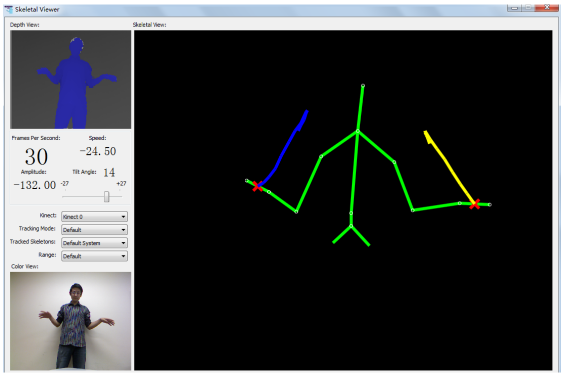


Fig. 6. GUI of Hand Conducting Gesture Recognition

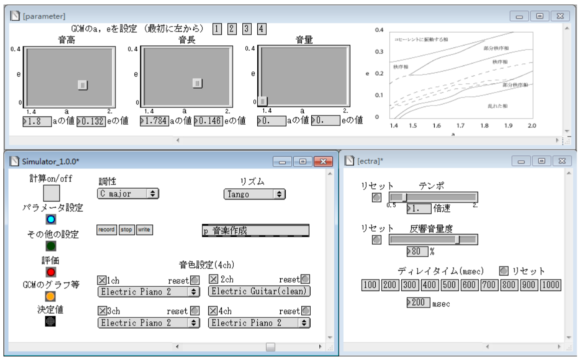


Fig. 7. GUI of ICAS Simulator

6. Simulation

We have implemented the system described in this paper to evaluate the effect of the system. We designed two experiments, to find out the difference of the output sound, when the input hand gesture have difference in speed and amplitude.

6.1. ICAS Simulator

The system we have simulated, is programmed under Microsoft Visual C++ and Cycling74 company's MAX/MSP. The GUI of hand conducting gesture recognition is shown in Fig. 6. We get the operator's music command gesture using a Kinect. The red cross shows the center point of hand and wrist. In this GUI, we calculate the hand speed and amplitude in real time, and we can get hand direction data easily, all these data will be send to the music generation and music arrangement program.

The GUI of improved ICAS simulator is shown in Fig. 7. This GUI mainly shows about the music generation, we can set the sound tonality and length here. And before our conducting, we can listen to the automatically generated music first, while adjusting the parameters. When we think it become interesting enough, we start our conducting, and will let the music richer, more diverse.

6.2. Simulation results

We performed an experiment aiming to verify the hand gesture and output sound associated well with each other. First, we disconnect the hand conducting parts, the output music will be automatic generated and without arrangement, we record the output music and get its score shown in Fig. 8. Second, we connect the hand conducting part and let the operator doing conducting gesture, and try to change the tempo and volume of the output music. Sample music score of second experiment is shown in Fig. 9.

In Fig. 8, the notes are stable and changes slowly. In Fig. 9, obviously different from Fig. 8, the generated notes are richer. At first, the operator conducts at a normal speed, and then waves hand fast, that the tempo changes quite match with the operator's hand conducting, in the first two section, the generated music at a normal speed, and then become fast in section three and four. From section five to section six, the tempo was back into leisurely and slowly.

As a result of the simulation, we confirmed that the generated music has more interactivity than the simple ICAS system and we can get the desired sound speed and volume roughly by changing our hand gestures, instead of adjusting the GCM parameters directly.

6.3. Kansei Evaluation

In this experiment, we try to find out that whether the music after arrangement by hand gesture can give listener a better impression than the music generated only by computer. In this research we use the Semantic Differential method to find out how the users think about our output music. Each subject was asked to hear about the result music of our system, and express his or her perceptions about the music in several adjectives, we record these words and find the antonym of them in the dictionary. We have prepared 10 pairs of bipolar adjectives with higher mentioned times to do this experiment, and we consider these 10 pairs of items have the same importance. The questionnaire is designed as shown in Tab. 1. Ten people around twenty years old, act as subjects in this experiment. We focus on the different feelings of subjects when hearing the output music which generated by ICAS only(open the track 1 and track 2 only) or generated by automatically rhythm and chord progression with ICAS synchronization, or generated by include all above and add the hand conducing arrange the music in real time. They were asked to hear these three outputs each for one minute, and to fill in the questionnaire. We try to find some differences during these questionnaire data. Hence, we have to avoid the interference caused by other variables. In this experiment, we set all the other variables as a constant. Here we set a=1.8 and e=0.13, and use the key C with tango rhythm pattern.

The questionnaire based evaluation result is shown in Fig. 10. In this figure, The vertical axis on the left side

Fig. 8. Score of Output Music (1)

Generated Music Without Gesture Arrangement
(C major, Tango)

Fig. 9. Score of Output Music (2)

Generated Music With Gesture Arrangement
(C major, Tango, Another Piece of Output Music)

Table 1. Questionnaire Using Semantic Differential

	Positive	3	2	1	0	1	2	3	Negative
D1	Strong								Feeble
D2	Bright								Dark
D3	Fantastic								Mechanical
D4	Intense								Calm
D5	Complexity								Monotone
D6	Impressive								Thin Impression
D7	Novelty								Common
D8	Good articulate								Bad articulate
D9	Light								Heavy
D10	Natural								Unnatural

shows ten Kansei adjectives that we considered to be positive, and the ten adjectives on the right vertical axis are considered to be negative. The horizontal axis indicates to the degree of the tendency between the two opposite sides. This tendency of feeling is divided into seven different levels, corresponding to the options in the questionnaire. The value of Fig. 10 is determined by the average value of questionnaire results.

The blue line in Fig. 10 shows the evaluation result of music generated by ICAS only, we can see the survey results indicated that the piece of music generated by ICAS only, was thought most monotone among the three, and all the subjects thought the output music of this generation method is more intense than the other two. The green line shows the result after the rhythm patterns and chord progression added into ICAS, which have the higher level of feeble and bright. And the red line shows the survey result of the output music after we arranged the music with hand gesture, and included all the method above, we found that the output music group of hand conducted was more complex and more impressive than the former ones. And the hand conducted generation was rated as a higher novelty. Both developed ICAS in this research have enriched the generated music, in the aspects of calm, complexity impressive and novelty. And lost some intense and concise.

Fig. 11 shows the range of deviation in the evaluation, the vertical axis is the ten pairs of adjectives of positive, and the horizontal axis is the level of the tendency. The deviation of ICAS only is the smallest, means that the subjects have more consensus on this kind of output, and these results have the apparently tendency on feeble, calm, monotone, thin impression, and heavy. While the deviation of ICAS+R+C and ICAS+R+C+Hand is larger. This might be because of several reasons. First, comparing with the output of ICAS only, the latter two methods are richer in variation, thus different people may get different feelings. Second, the latter two output are generated with more music elements, and maybe more similar with music, as a result, it can be easier to fell variable emotions in the music. This can find obviously on the red deviation line of fantastic, subjects' opinion about fantastic is quite different, from -0.5 somewhat not fantastic, to nearly +2 general fantastic.

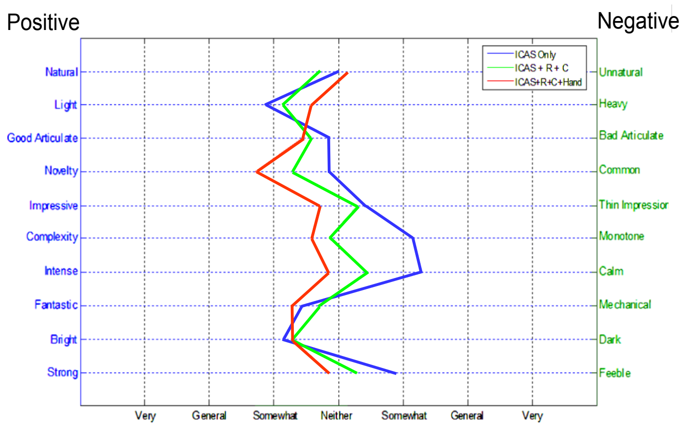


Fig. 10. Kansei Evaluation Results

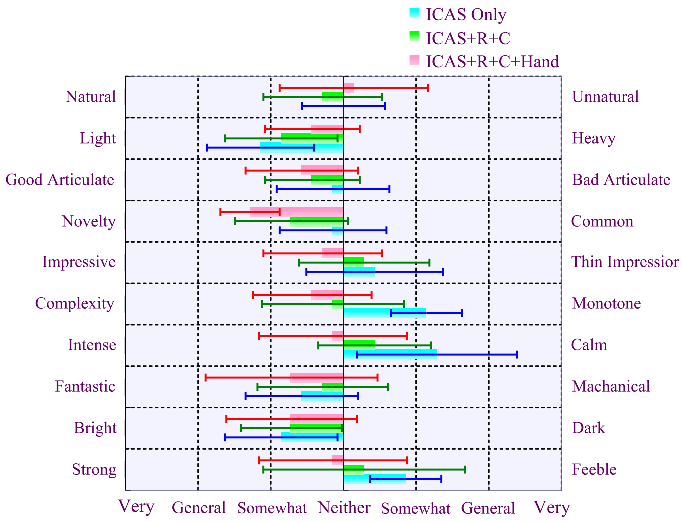


Fig. 11. Kansei Evaluation Deviation

7. Conclusions

In this paper, we proposed the music conductor gesture arranged music generation system. We generated the chaotic sound first, the conducting hand gestures have been used to arranging the output music. We took the rhythm and chord progression of music into consideration. On the other hand, the conducting gesture was also considered the simplest situation of 2/2 or 2/4 time. Kansei evaluation have been done, we think the result have showed that the output music group of hand conducted was more complex and more impressive than the former ones.

For the further research, we will enrich the musical expressions and the conducting gestures to achieve a better output sound and a higher level of interaction.

References:

[1] A. Mulder, S. Fels, "Sound Sculpting: Manipulating Sound through Virtual Sculpting", *Western Computer Graphics Symposium*, pp. 15-23, 1998.

[2] H. S. Ip, H. Young, C. C. Tang, "Body-Brush: A Bodydriven Interface for Visual Aesthetics", *ACM International Conference on Multimedia 2002*, pp. 664-665, 2002.

[3] H. S. Ip, C. K. Law, B. Kwong, "Cyber Composer: Hand Gesture-Driven Intelligent Music Composition and Generation", *Proceed-*

ings of the 11th International Multimedia Modeling Conference (MMM '05), Vol. 1550, pp. 5502-5505, 2005.

[4] M. Yamamoto, Y. Maeda, "Interactive Kansei System for Sound Generation Based on Network of Chaotic Elements ", *Proc. of the Fourth International Symposium on Human and Artificial Intelligence Systems: From Control to Autonomy (HART 2004)*, pp. 315-320, 2004.

[5] K. Kaneko, "Partition complexity in a network of chaotic elements", *Jounal of Physics A*, pp. 2107-2119, 1991.

[6] B. Benward, M. Saker, "Music: In Theory and Practice", p.213, 2008.

[7] A. Blatter, "Revisiting music theory: a guide to the practice", p.28, 2007.

Name:
Shuai Chen

Affiliation:
Department of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui

Address:
3-9-1 Bunkyo, Fukui 910-8507 Japan

Brief Biographical History:
2007- Major of Automation, School of Information Engineering, Nanchang Hangkong University
2012- Dept. of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui

Membership in Learned Societies:

Name:
Yoichiro Maeda

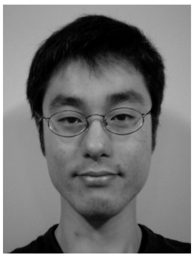
Affiliation:
Department of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui

Address:
3-9-1 Bunkyo, Fukui 910-8507 Japan

Brief Biographical History:
1983- Researcher, Central Research Lab., Mitsubishi Electric Corp.
1989-1992 Senior Researcher, Laboratory for International Fuzzy Engineering Research (LIFE)
1995- Associate Prof., Osaka Electro-Communication University
1999-2000 Visiting Researcher, University of British Columbia (UBC), Canada
2002- Associate Prof., Faculty of Engineering, University of Fukui
2007- Prof., Graduate School of Engineering, University of Fukui

Main Works:
• Y. Maeda, M. Tanabe, and T. Takagi, "Behavior-Decision Fuzzy Algorithm for Autonomous Mobile Robots," Information Sciences, Vol.71, No.1, pp.145-168, 1993.
• Y. Maeda, "Emotional Generation Model for Autonomous Mobile Robot," KANSEI Engineering Int., Vol.1, No.1, pp.59-66, 1999.
• R. Taki, Y. Maeda, and Y. Takahashi, "Personal Preference Analysis for Emotional Behavior Response of Autonomous Robot in Interactive Emotion Communication," Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII), Vol.14, No.7, pp.852-859, 2010.

Membership in Learned Societies:
• The Society of Instrument and Control Engineers (SICE)
• The Robotics Society of Japan (RSJ)
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
• The Japanese Society for Artificial Intelligence (JSAI)
• Japan Society of Kansei Engineering (JSKE)



Name:
Yasutake Takahashi

Affiliation:
Department of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui

Address:
3-9-1, Bunkyo, Fukui, 910-8507, Japan

Brief Biographical History:
2000-2009 Assistant Professor of Department of Adaptive Machine Systems, Graduate School of Engineering, Osaka University
2003-2009 Member of exec committee for RoboCup middle size league
2006-2007 Visiting researcher at the Fraunhofer IAIS
2009- Senior Assistant Professor of Department of Human and Artificial Intelligent Systems, Graduate School of Engineering, University of Fukui

Main Works:
• Yasutake Takahashi, Kentaro Noma, and Minoru Asada. Efficient Behavior Learning based on State Value Estimation of Self and Others. Advanced Robotics, Vol.22, No.12, pp.1379-1395, 2008.
• Yasutake Takahashi, Yoshihiro Tamura, Minoru Asada, and Mario Negrello. Emulation and behavior understanding through shared values. Robotics and Autonomous Systems, Vol.58, No.7, pp.855–865, 2010.
• Yoshihiro Tamura, Yasutake Takahashi, and Minoru Asada. Observed Body Clustering for Imitation Based on Value System. Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.14, No.7, pp.802–812, 2010.

Membership in Learned Societies:
• The Robotics Society of Japan (RSJ)
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
• The Japanese Society for Artificial Intelligence (JSAI)
• The Institute of Electrical and Electronics Engineers (IEEE)